



Une mesure économétrique de la formation par les autres au sein de l'entreprise

Guillaume Destré

► To cite this version:

Guillaume Destré. Une mesure économétrique de la formation par les autres au sein de l'entreprise. 2005. halshs-00193950

HAL Id: halshs-00193950

<https://shs.hal.science/halshs-00193950>

Submitted on 5 Dec 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Une mesure économétrique de la formation
par les autres au sein de l'entreprise**

Guillaume DESTRE, TEAM

2005.21

Une mesure économétrique de la formation par les autres au sein de l'entreprise[^]

Guillaume Destré[♦]

[^] Je remercie Andrew Clark, Louis Lévy-Garboua, David Margolis et Michel Sollogoub pour leurs remarques constructives.

[♦] TEAM, 106-112 Bd de l'Hôpital, 75 647 Paris Cedex 13.
E-mail: gdestre@univ-paris1.fr

A partir de l'estimation de la fonction de gains nets d'un modèle d'apprentissage qui identifie une des principales composantes de la formation en entreprise: le learning by watching, nous proposons, à l'aide de données appariées de l'INSEE sur le coût de la main-d'œuvre et la structure des salaires en 1992 (ECMOSS 92), une mesure économétrique de la formation sur le tas. Cette dernière s'élève à 42,75 heures en moyenne par an et représente près de 80% de la formation totale dispensée en entreprise. La richesse du modèle présenté ici ne réside pas exclusivement dans la mesure de la formation sur le tas mais aussi, dans l'estimation respective de la vitesse à laquelle se diffuse le savoir au sein de l'établissement et du savoir de l'entreprise relativement à celui du travailleur entrant.

Mots clés: Capital humain, fonctions de gains, formation informelle, apprentissage par observation des autres.

We use a model of the diffusion of knowledge which considers one of the principal forms of training found within firms: learning by watching. Using a French INSEE matched employee-employer survey (le coût de la main-d'œuvre et la structure des salaires 1992: ECMOSS 92), we propose an econometric measure of on-the-job training. The results show that on-the-job training amounts to 42,75 hours in the mean by year and represents nearly 80% of the total training in the firm. The richness of the model presented here consists in the estimation not only of on-the-job-training, but also of the speed of knowledge diffusion in an establishment and the relative knowledge of the firm (relative to that of a new worker).

Keywords: Human capital, earnings functions, informal training, learning from others.

Classification JEL: J24, J31.

INTRODUCTION

Depuis les travaux de Becker (1962) et Mincer (1962), la théorie du capital humain considère que la croissance des salaires avec l'ancienneté est le résultat de la formation sur le tas (*on-the-job-training*). Pourtant, les économistes éprouvent les pires difficultés à vérifier cette assertion tant il est difficile de mesurer certains types d'investissement en capital humain comme la formation informelle. En effet, comme le remarque fort justement (Brown, 1989):

"While there are difficulties in measuring formal training, what we would like to measure is relatively well-defined: an individual is either in a training program or not, formal training has an identifiable start and end, and one should in principle be able to determine either how many hours the worker spent or how many dollars the employer spent on any particular training.

In contrast, informal training is produced jointly with the primary output of the worker, and is therefore more elusive. Workers learn from watching other workers, may share easier ways to do the work either while working or during breaks, and are indirectly instructed whenever a supervisor constructively criticizes their work. Knowing whether informal training is happening in any given week is difficult to determine; one hopes that for most workers it never ends. The dollar cost is elusive not only because the time spent by supervisors and other workers is not logged, but also because the worker's productivity is also likely to be reduced while in training".

Dans ces conditions, il n'est guère surprenant que, d'une part les quelques tentatives visant à mesurer la formation dispensée par les employeurs se soient concentrées essentiellement sur la formation formelle et que, d'autre part, les mesures directes de la formation informelle dont nous disposons soient très disparates et entachées d'erreurs, ce qui conduit notamment à biaiser à la baisse les taux de rendement sur les gains (Barron *et al.* 1997a et 1997b et Loewenstein et Spletzer, 1998).

Selon l'OCDE (1997), la formation informelle dispensée par les employeurs semble avoir joué un rôle important dans la formation à des emplois se situant au bas de la hiérarchie professionnelle, dont un grand nombre de compétences peuvent être acquises à peu de frais sur le tas et dans la formation du personnel des organisations de petite taille où la formation formelle coûte trop cher en termes de journées de travail perdues, du fait qu'elle est, pour l'essentiel assurée en dehors du lieu de travail. Toutefois, on observe actuellement que la formation informelle prend une importance croissante dans la mesure où elle joue un rôle de plus en plus grand pour des emplois se situant dans la partie supérieure de la hiérarchie professionnelle et pour des organisations de grande taille. Ce phénomène tient sans doute à

l'accroissement des besoins de formation lié au vaste changement technologique et structurel en cours qui rend de plus en plus coûteux pour les employeurs le recours à la formation formelle.

Ainsi, il ne suffit plus de mesurer la formation formelle dispensée par les employeurs pour évaluer de manière adéquate l'ensemble de la formation qu'ils organisent. Toujours selon l'OCDE (1997), l'absence de recensement de la formation informelle conduit non seulement à une évaluation incorrecte de la formation nationale, mais fausse également les comparaisons entre pays ou secteurs dès lors que ceux-ci ont recours aux formations de type formel ou informel dans des proportions différentes.

L'objet de cet article sera de proposer une mesure économétrique de la formation informelle qui soit susceptible de s'affranchir des erreurs de mesure propres aux enquêtes. Celle-ci s'appuiera sur l'estimation de la fonction de gains nets d'un modèle d'apprentissage¹ qui identifie une des principales composantes de la formation informelle: le *learning by watching* (Lévy-Garboua [1994])² et sur les données appariées de l'enquête française de l'INSEE sur le coût de la main-d'œuvre et la structure des salaires en 1992 (ECMOSS 92). Le volume de la formation informelle dispensée au sein des établissements pourra alors être comparé à celui de la formation formelle, directement observé par cette même enquête.

Les enjeux sont d'une part, la possibilité pour les pouvoirs publics d'avoir une vue complète du système national de formation en déterminant ce que représente la formation informelle dans la formation totale (notamment en vérifiant l'existence ou non d'une complémentarité ou d'une substituabilité entre les deux types de formation que sont la formation formelle et la formation informelle), ce qui devrait les aider à élaborer une politique de formation efficace et d'autre part, la possibilité d'effectuer des comparaisons internationales

¹ La fonction de gains brut de ce modèle d'apprentissage a déjà été estimée à plusieurs reprises, successivement par Chennouf, Lévy-Garboua et Montmarquette [1997], Nordman [2000] et Destré [2003] à l'aide d'enquêtes employeurs-employés (canadienne et algérienne pour les premiers, marocaine et mauricienne pour le second et française pour le troisième) puisque la modélisation d'un processus d'apprentissage informel susceptible d'être soumis à l'épreuve des faits exige la disponibilité de données microéconomiques contenant des informations à la fois sur les individus mais aussi sur les entreprises au sein desquelles ces derniers évoluent.

² Les résultats d'enquêtes décrivant les pratiques des entreprises américaines au cours des premiers mois suivant l'embauche indiquent en effet qu'à peu près 90% de la formation sur le tas est dispensée de manière informelle par le personnel de l'entreprise et qu'environ le tiers de la formation sur le tas aurait effectivement lieu par *learning by watching* (Barron, Berger et Black [1997a] et [1997b]).

et intranationales puisqu'en étant issue des profils de gains individuels, la mesure de la formation proposée ici est parfaitement homogène.

Après avoir explicité le modèle dans une première partie et présenté les données (l'enquête appariée de l'INSEE sur le coût de la main-d'œuvre et la structure des salaires en 1992) et les méthodes économétriques dans les deuxième et troisième parties, nous analyserons les principaux résultats dans la quatrième partie avant de conclure.

LE MODÈLE³

Dans ce modèle, le processus de formation s'effectue par une sorte de *learning by watching* (LBW), où chaque salarié apprend de tous ceux au savoir desquels il est exposé et qui en savent plus que lui. Ce mécanisme d'apprentissage ne "gaspille" pas le temps des enseignants mais seulement celui des apprentis⁴.

La production du capital humain au sein de l'entreprise

Le capital humain est produit par l'effort de formation. Comme on suppose qu'il n'y a pas de dépréciation⁵, on peut écrire la fonction de production du capital humain:

$$(1) \quad h_t - h_{t-1} = h_{t-1} r e_t$$

avec:

$$(2) \quad R_t = r e_t$$

où h_{t-1} représente le capital humain d'un individu quelconque à la fin de la période $t-1$, r est le taux de rendement moyen de la formation (supposé constant conformément à Mincer [1974]), e_t désigne le temps de formation au cours de la période t avec $0 \leq e_t \leq T$, T étant le temps

³ Pour une présentation plus détaillée du modèle, on pourra se référer à Lévy-Garboua [1994].

⁴ Le LBW étudié ici est donc un LBW au sens large dans la mesure où il recouvre tous les types de formation informelle systématique dans l'entreprise pour lesquels il n'y a pas de coût pour le formateur. Par conséquent, il ne se limite pas à la simple observation des autres.

⁵ Ce qui est une bonne approximation pour les salariés qui n'interrompent pas leur activité.

marchand et R_t se définit comme le rendement marginal brut de l'ancienneté au cours de la période t .

On suppose que la formation reçue par salarié au cours de la période t , après $t-1$ années passées dans l'entreprise, est proportionnelle au différentiel de connaissances entre lui (l'élève) et celui qui en sait le plus parmi ceux au savoir desquels il est en contact (l'enseignant):

$$(3) \quad h_t - h_{t-1} = \frac{n}{1+n} (H - h_{t-1})$$

où H désigne le savoir de l'entreprise (incorporé au sein de celui qui en connaît le plus puisque le savoir est supposé homogène) et n est un paramètre positif (réputé invariant dans le temps, pour simplifier), qui résume le rythme de diffusion du savoir des enseignants à leurs élèves. Un simple réarrangement des termes montre que le capital humain est une moyenne pondérée entre le capital humain en $t-1$ et le savoir maximal H auquel l'individu se trouve exposé dans l'entreprise:

$$(4) \quad h_t = \frac{1}{1+n} h_{t-1} + \frac{n}{1+n} H$$

Si l'on suppose, dans un marché concurrentiel, que les taux de salaire sont égaux à la productivité marginale des individus, les gains bruts représentent à chaque instant le capital humain des salariés. A partir de l'équation (3), un calcul par récurrence donne la fonction de gains bruts suivante (puisque'on ne défalque pas le coût du temps de la formation supporté par "l'apprenti"):

$$(5) \quad h_t = k^t h_0 + (1 - k^t) H \quad \text{avec} \quad k = \frac{1}{1+n}$$

où le savoir du salarié est une moyenne pondérée de son savoir initial, au moment d'entrer dans l'entreprise, et du savoir maximal auquel il peut accéder dans cette entreprise. Le savoir individuel converge vers le savoir "social" auquel l'individu est exposé selon une fonction exponentielle de l'ancienneté.

En réalité, ni les gains bruts ni le gains nets ne sont réellement observés. Cependant, Mincer [1974] considère que les gains observés s'apparentent davantage aux gains nets qu'aux gains bruts. De plus, la fonction de gains bruts ne nous permet pas de mesurer la formation sur le tas puisque ni r_t ni e_t ne sont identifiés. Il nous faut donc réécrire la FGB en tenant compte du coût de la formation supporté par le salarié.

Une fonction de gains nets

La formation étant informelle, nous pouvons considérer que le coût de la formation est essentiellement un coût d'opportunité du temps. Les gains nets (w_t) s'obtiennent alors à partir des gains bruts (h_t) en soustrayant le coût du temps de formation (Becker [1975] et Mincer [1974]):

$$(6) \quad w_t = h_t (T - e_t)$$

A partir de l'équation (3), nous pouvons calculer le taux de rendement marginal brut de la $t^{ième}$ période d'ancienneté (pour $t \geq 1$):

$$(7) \quad R_t = \frac{h_t - h_{t-1}}{h_{t-1}} = \frac{n\lambda k^t}{\lambda(1 - k^{t-1}) + 1}$$

avec $\lambda = \left(\frac{H}{h_0} - 1 \right)$ où $\lambda \geq 0$ désigne le potentiel d'apprentissage du travailleur dans son entreprise.

En remplaçant e_t par $\frac{R_t}{r}$ (voir (2)) et h_t par (5), nous obtenons l'expression suivante pour $T \equiv 1$ et $t \geq 1$:

$$(8) \quad w_t = \left[k^t h_0 + (1 - k^t) H \right] \left[1 - \frac{n}{r} \left(\frac{\lambda k^t}{(\lambda + 1) - \lambda k^{t-1}} \right) \right]$$

Une mesure de la formation sur le tas

L'estimation de r et le calcul de R_t permettent de calculer le temps de formation sur le tas de chaque individu à la date t puisque d'après (2):

$$(9) \quad e_t = \frac{R_t}{r} = \frac{\frac{n\lambda k^t}{(\lambda + 1) - \lambda k^{t-1}}}{r}$$

La possibilité conférée à chaque salarié d'apprendre de ses collègues a pour effet de rendre la formation sur le tas dépendante de la firme qui l'emploie. Cette dernière est, plus précisément, une fonction croissante concave du potentiel d'apprentissage du travailleur dans son entreprise:

$$\frac{\partial e_t}{\partial \lambda} > 0 \qquad \frac{\partial^2 e_t}{\partial \lambda^2} < 0$$

L'expression (9) indique également que la formation sur le tas est une fonction décroissante convexe de l'ancienneté, alors que l'hypothèse, prônée par Mincer et Jovanovic [1981] d'une fonction de gains quadratique impliquait une relation linéaire.

$$\frac{\partial e_t}{\partial t} < 0 \qquad \frac{\partial^2 e_t}{\partial t^2} > 0$$

LES DONNÉES

Comme nous l'avons souligné en introduction, la modélisation d'un processus d'apprentissage informel exige la disponibilité de données microéconomiques contenant des informations à la fois sur les individus et sur les entreprises au sein desquelles ces derniers

évoluent. L'enquête française de l'INSEE sur le coût de la main d'œuvre et la structure des salaires en 1992 (ECMOSS 92) appartient à cette catégorie de données puisqu'elle croise un échantillon de 150 000 salariés avec un échantillon de 16 000 établissements (voir Araï, Ballot et Skalli [1996] et Abowd, Kramarz, Margolis et Troske [2000] entre autres).

Pour pouvoir procéder aux estimations, nous avons dû construire un certain nombre de variables parmi lesquelles: le nombre d'années d'études, l'expérience potentielle sur le marché du travail (âge – nombre d'années d'études – six), le gain horaire ((rémunération brute totale + avantages en nature) / nombre d'heures rémunérées dans l'année), le nombre moyen d'heures de formation rémunérées par salarié dans l'établissement au cours de l'exercice (nombre d'heures de formation rémunérées par catégorie sociale (cadre ou non-cadre) / effectif salarié sous contrat de la catégorie sociale considérée).

Concernant l'éducation, un nombre d'années d'études théorique a été reconstitué pour chaque individu à partir du diplôme le plus élevé déclaré pour lui (par l'établissement). Pour cela, à l'aide d'un échantillon d'un peu plus de 8 000 salariés (réalisé conjointement à ECMOSS 92) pour lequel l'année de fin d'études est renseignée en plus du diplôme⁶, nous avons calculé le nombre d'années d'études médian (moins sensible aux points aberrants que la moyenne dans le cas de petits échantillons) correspondant à chacun des diplômes⁷.

Après élimination des valeurs manquantes et aberrantes, l'échantillon sur lequel porte l'étude s'élève à 119 667 individus répartis dans 14 449 établissements. Le tableau 1 de l'annexe A résume les symboles et définitions des principales variables ainsi que les statistiques descriptives.

⁶ En effet, en plus de la partie salariés d'ECMOSS 92 (renseignée par les établissements), un questionnaire personnel a été envoyé directement à un sous-échantillon de salariés.

⁷ L'enquête ECMOSS 92 comporte un nombre important d'individus pour lesquels le diplôme le plus élevé est non renseigné (environ un tiers de l'échantillon). Plutôt que d'éliminer ces observations, ce qui aurait conduit à biaiser les résultats, nous avons choisi de préserver la représentativité de l'échantillon et d'affecter également à cette catégorie d'individus le nombre d'années d'études médian leur correspondant dans le sous-échantillon de 8000 salariés. Bien entendu, cette façon de procéder peut être discutée puisque le nombre élevé de salariés concernés par cette situation suggère une importante hétérogénéité des parcours scolaires. Cela étant, comme l'a montré Destré [2003], le fait de supprimer les observations pour lesquelles le diplôme est non renseigné n'entraîne pas de modifications substantielles de l'ensemble des résultats.

LES MÉTHODES ÉCONOMÉTRIQUES

Le savoir de l'établissement H , de même que le savoir individuel h_0 , ne sont pas observables. Nous les approximations l'un et l'autre par une fonction de Mincer [1974] quadratique dans les termes d'expérience et d'ancienneté. L'estimation du salaire à l'entrée dans le dernier établissement h_0 ne comporte évidemment pas de termes d'ancienneté (dans ce même établissement), et s'écrit:

$$(10) \quad h_{i0} = \exp(a_0 + a_1 s_i + a_2 x_i + a_3 x_i^2)$$

où s_i et x_i désignent respectivement le nombre d'années d'études et d'expérience potentielle de l'individu i hors de l'établissement j (c'est-à-dire avant d'y entrer).

Par ailleurs, nous supposons que le savoir de l'établissement auquel l'individu i est exposé est incorporé dans un individu – fictif – ayant S_{ij} années d'éducation, X_{ij} années d'expérience hors de l'établissement, et T_{ij} années d'ancienneté dans l'entreprise actuelle, tels que $S_{ij} \geq s_i$, $X_{ij} \geq x_i$, $T_{ij} \geq 0$. Il est estimé, à son tour, par la fonction de Mincer [1974] quadratique:

$$(11) \quad H_{ij} = \exp(a_0 + a_1 S_{ij} + a_2 X_{ij} + a_3 X_{ij}^2 + a_4 T_{ij} + a_5 T_{ij}^2)$$

S_{ij} , X_{ij} , T_{ij} représentent respectivement le maximum de S , X et T auquel i est exposé dans j . Il convient de préciser que chaque salarié, dans notre modèle, a un savoir différent qui dépend de son propre niveau d'éducation, d'expérience et d'ancienneté⁸. Ces caractéristiques ne sont pas disponibles dans nos données⁹. Cependant, nous observons ces variables pour un échantillon aléatoire d'individus d'un même établissement. Une façon pratique de retrouver les caractéristiques inconnues (éducation, expérience, ancienneté) de l'enseignant le plus qualifié du travailleur i est de supposer que la valeur prise par chacune d'elles se situe entre le maximum observable au sein de l'échantillon de l'établissement j et celle de l'individu i . Si $z_{ij} = (s_i, x_i, t_{ij})$ et $Z_{ij} = (S_{ij}, X_{ij}, T_{ij})$ représentent les vecteurs de capital humain de l'individu i et de son enseignant dans l'établissement j et $Z_j = \sup_{i \in j} z_{ij}$ est la valeur maximum observable de la

⁸ H sera non seulement indicé par j mais aussi par i .

⁹ De telles observations nécessiteraient de connaître les interactions entre les individus au sein des établissements.

caractéristique z dans le sous-échantillon de l'établissement, nous pouvons écrire Z_{ij} comme une moyenne pondérée de Z_j et z_{ij} :

$$(12) \quad Z_{ij} = \beta_z Z_j + (1 - \beta_z) z_{ij} \quad \text{avec} \quad 0 \leq \beta_z \leq 1$$

β_z prend la valeur zéro si l'individu n'a pas de possibilité d'apprentissage par les autres et la valeur 1 si son enseignant le plus qualifié correspond effectivement au travailleur ayant le z le plus élevé au sein de l'échantillon de l'établissement¹⁰.

Après avoir reporté (12) dans (11), nous pouvons utiliser (10) pour calculer:

$$(13) \quad \lambda_{ij} = \exp \left(a_1 \beta_s (S_j - s_i) + a_2 \beta_x (X_j - x_i) + a_3 \beta_x^2 (X_j - x_i)^2 + 2a_3 \beta_x (X_j - x_i) x_i + a_4 \beta_t T_j + a_5 \beta_t^2 T_j^2 \right) - 1$$

En reportant (13) et (10) dans l'équation (8) (après avoir mis en facteur h_0), nous obtenons (après passage aux logarithmes) pour l'individu i dans l'établissement j une fonction de gains nets non linéaire à laquelle nous pouvons ajouter un terme aléatoire u_{ijt} :

$$(14) \quad \ln w_{ijt} = \ln h_{i0} + \ln \left[1 + \lambda_{ij} (1 - k^{t_{ij}}) \right] + \ln \left[1 - \frac{n}{r} \left(\frac{\lambda_{ij} k^{t_{ij}}}{1 + \lambda_{ij} (1 - k^{t_{ij}-1})} \right) \right] + \sum_k \delta_k c_k + u_{ijt}$$

où c_k représente un ensemble de variables de contrôle, et δ_k les coefficients associés à chacune de ces variables. Nous estimerons cette équation à l'aide des moindres carrés non linéaires (MCNL).

Le salaire de tout individu se décompose log-additivement en son salaire à l'entrée dans sa dernière entreprise et en deux termes qui ne dépendent que du savoir relatif de celle-ci

¹⁰ Afin de tenir compte du fait que, pour chacune des variables de capital humain, le maximum observé peut être dans certains cas inférieur au maximum auquel les individus sont réellement exposés (puisque nous travaillons sur un échantillon et non sur la population), nous ne contraindrons pas les paramètres β_z à être compris entre 0 et 1 dans l'estimation. Ainsi, s'il s'avère que le maximum observé est inférieur au maximum réel, β_z pourra prendre une valeur supérieure à 1 et l'estimation de H/h_0 ne sera pas affectée par la nature de nos données. En revanche, il conviendra de prendre en considération le caractère très flexible de notre forme fonctionnelle dans l'interprétation des différents β_z estimés puisque ces derniers seront probablement surévalués.

par rapport au sien, et de sa propre ancienneté. En pratique, l'entreprise est remplacée par l'établissement qui est l'unité enquêtée, ce qui semble être un bon compromis empirique puisque les salariés sont évidemment, plus exposés aux autres salariés du même établissement qu'à toute autre personne.

En l'absence de dimension temporelle, l'hétérogénéité inobservée des individus ne peut être prise en compte. En revanche, l'hétérogénéité des établissements peut être traitée, du moins en théorie (Abowd *et al.* [2000]). L'inconvénient est que nous sommes ici en présence d'un modèle non-linéaire. Deux solutions pourraient permettre de régler le problème. La première consiste à introduire un nombre de variables muettes égal au nombre d'établissements. Malheureusement, eu égard au nombre d'établissements présents dans l'enquête, cette proposition n'est pas envisageable. La seconde, qui consiste à linéariser le modèle et à appliquer les techniques économétriques usuelles, s'avère difficile à mettre en œuvre sans recourir à des hypothèses fortes (pour la linéarisation). C'est pourquoi, afin de tempérer les effets de l'hétérogénéité des établissements et des individus (qui peut biaiser les coefficients de régression), nous avons introduit un nombre élevé de variables de contrôle¹¹.

¹¹ Les variables de contrôle retenues pour cette étude sont (la catégorie de référence figure entre parenthèses): le logarithme du nombre d'heures de travail, une indicatrice du sexe (femme), trois indicatrices de la nationalité (français), cinq indicatrices de la situation familiale (marié), trois indicatrices du type de contrat de travail (CDI), une indicatrice de la CSP (non-cadre), une indicatrice de la région (province), une indicatrice de la présence de syndicats (absence de Délégués Syndicaux), trois indicatrices de l'existence d'un système formalisé d'évaluation des performances individuelles (existence d'un tel système), quatre indicatrices de l'incitation à la coopération directe entre salariés du même service (existence d'une telle incitation), trois indicatrices de l'existence de rotations de postes (absence de rotations de postes), cinq indicatrices prenant en compte si l'évolution des salaires de base dépend [exclusivement, principalement, peu, pas du tout, non renseignée] des augmentations individualisées (peu), cinq indicatrices prenant en compte l'importance accordée [aucune, faible, moyenne, très forte, non renseignée] à l'ancienneté dans la progression individualisée du salaire (aucune), cinq indicatrices prenant en compte l'importance accordée [aucune, faible, moyenne, très forte, non renseignée] à l'accroissement des performances du salarié dans la progression individualisée du salaire (très forte), cinq indicatrices prenant en compte l'importance accordée [aucune, faible, moyenne, très forte, non renseignée] à l'effort de formation du salarié dans la progression individualisée du salaire (moyenne), cinq indicatrices prenant en compte l'importance accordée [aucune, faible, moyenne, très forte, non renseignée] à l'accumulation d'expérience du salarié dans la progression individualisée du salaire (moyenne), quatre indicatrices du nombre d'heures moyen de formation formelle rémunérées par l'établissement au cours de l'année (aucune heure de formation), douze indicatrices de secteur (services marchands) et six indicatrices de la taille de l'établissement (moins de 20 salariés).

LES PRINCIPAUX RÉSULTATS

Les fonctions de gains

Le tableau 2 de l'annexe A présente les résultats des différentes fonctions de gains. Alors que la première colonne expose les estimations obtenues à partir d'une fonction de gains mincerienne (FGM) standard (quadratique en ancienneté), les colonnes 2 et 3 répertorient respectivement les estimations des fonctions de gains bruts (FGB)¹² et nets (FGN) de notre modèle d'apprentissage par LBW.

Tous les coefficients et paramètres sont statistiquement significatifs au seuil de 1%. L'estimation de la fonction de gains nets (14) (colonne 3) accroît de façon significative le pouvoir explicatif de notre modèle puisqu'un test de Fisher du modèle contraint (FGB) contre le modèle non-contraint (FGN) conduit à rejeter l'hypothèse de base $H_0: \rho=0$ (avec $\rho=1/r$) au seuil de 1%. La non-prise en compte du coût d'opportunité de la formation dans la fonction de gains de bruts conduit à sous-estimer la vitesse de diffusion du savoir puisque celle-ci passe de 4,25% (colonne 2) à 5,49% (colonne 3). Cette différence est encore plus saillante lorsqu'on compare le temps nécessaire à un individu pour acquérir la fraction α du savoir de l'établissement au sein duquel il travaille. On dérive en effet de l'équation (5):

$$(1+n)^t = \frac{H-h_0}{H-h_t} = \frac{1}{\alpha}$$

Le tableau 3 de l'annexe A montre les résultats obtenus, en résolvant pour t , avec $\alpha=50\%$. Alors que la fonction de gains bruts stipule qu'il faut en moyenne plus de 16 ans à un individu pour apprendre 50% du savoir de son établissement, la fonction de gains nets montre qu'il n'en faut qu'un peu moins de 13. En revanche, la FGB surestime les possibilités d'apprentissage des individus au moment où ils entrent dans l'établissement puisque le savoir relatif (H_{ij}/h_{i0}) s'élève à 1,50 contre 1,40 avec la FGN (tableau 3 de l'annexe A). Cela signifie qu'un individu peut, en moyenne, apprendre l'équivalent de 50% de son capital humain initial dans le premier cas contre seulement 40% dans le second cas puisque $H_{ij}=1,40h_{i0}$.

¹² La FGB est définie par l'expression (14) avec $r=+\infty$.

La colonne 3 du tableau 2 indique également que la distance relative qui sépare le salarié moyen de son enseignant le plus qualifié est égale à 0,38 pour l'éducation (β_s) et 0,88 pour l'expérience initiale (β_x). Cette distance relative prendrait la valeur 0 s'il n'y avait aucune possibilité d'apprentissage pour le salarié et 1 si l'enseignant le plus qualifié coïncidait toujours avec le travailleur le plus qualifié de l'établissement dans l'échantillon. Bien que ces paramètres soient probablement surévalués (voir la note de bas de page numéro 10), il est intéressant de constater qu'ils sont strictement compris entre 0 et 1 alors que l'estimation a été réalisée sans contraintes. Là encore, lorsqu'on observe les colonnes 2 et 3 du tableau 2, on s'aperçoit qu'il existe quelques différences entre les résultats obtenus à partir de la FGB et ceux de la FGN. En effet, la fonction de gains bruts a tendance à surestimer, en moyenne, la distance relative séparant l'individu de son enseignant le plus qualifié. Cela étant, ces différences sont minimales et on ne commet pas une erreur importante lorsqu'on estime la FGB à la place de la FGN.

Enfin, lorsqu'on compare les rendements marginaux de l'éducation et de l'expérience initiale (Tableau 4 de l'annexe A), on constate que, quelle que soit la fonction utilisée (FGB ou FGN), ils sont du même ordre de grandeur que ceux obtenus par la FGM, ce qui suggère que cette dernière est une bonne approximation du savoir de base (c'est-à-dire à l'entrée dans l'établissement)¹³. Notons également que les résultats obtenus ici avec la FGB pour $t \geq 1$ sont quasiment identiques à ceux qu'avaient obtenus Destré [2003] avec cette même fonction pour l'ensemble de l'échantillon¹⁴. Ce résultat s'explique par le fait que les individus ayant une ancienneté inférieure à 1 an ne représentent que 12,7% de l'échantillon. Ainsi, tout en étant moins bien spécifiée que la FGN, la FGB apparaît suffisante (avec nos données) pour estimer les rendements de l'éducation et de l'expérience initiale.

La formation sur le tas

Le tableau 5 de l'annexe A permet de comparer les rendements marginaux bruts de l'ancienneté calculés à partir des trois fonctions de gains précédemment estimés. On constate que le rendement marginal brut de l'ancienneté, calculé à partir de (7) au point moyen de

¹³ Ce résultat avait déjà été observé par Chennouf *et al.* [1997], Nordman [2000] et Destré [2003] à partir de la FGB.

¹⁴ Puisque les taux de rendement de l'éducation et de l'expérience initiale calculés avec la FGN sont quasiment identiques à ceux calculés par Destré [2003] avec la FGB pour l'ensemble de l'échantillon, la comparaison réalisée par ce dernier avec les résultats obtenus par Goux et Maurin [1994] demeure valide.

l'ancienneté, s'élève à 1,06% avec la FGN. Ce rendement est significativement positif et inférieur à celui obtenu avec la FGM (test de Student d'égalité des moyennes). Ce résultat avait déjà été observé par Chennouf *et al.* [1997], Nordman [2000] et Destré [2003] à l'aide de la FGB. Notons que, comme pour les rendements de l'éducation et de l'expérience initiale, quelle que soit la fonction utilisée (FGB ou FGN), les taux de rendement de l'ancienneté sont du même ordre de grandeur. En revanche, lorsqu'on compare les résultats obtenus ici avec la FGB pour $t \geq 1$ avec ceux obtenus par Destré [2003] à l'aide de la même fonction pour l'ensemble de l'échantillon, on constate que ces derniers sont significativement plus élevés. Ce résultat s'explique par le fait que les rendements marginaux bruts de l'ancienneté sont une fonction décroissante de l'ancienneté (Lévy-Garboua [1994] et Destré [2003]).

L'estimation de r et le calcul de R_{ijt} permettent de déterminer, à l'aide de l'équation (9), la proportion de temps marchand (e_{ijt}) que les individus ont consacrée à la formation informelle au cours de l'année 1992. Lorsqu'on multiplie cette proportion par le nombre d'heures rémunérées au cours de l'année ($nbhran_i$), on obtient la durée de formation informelle. Celle-ci est décroissante avec le temps et s'élève à 47,38 heures au point moyen de l'ancienneté (tableau 6 de l'annexe A).

Le tableau 7 montre que le nombre d'heures moyen de formation informelle par salarié dans l'établissement est significativement plus élevé (42,75 heures contre 11,92) que celui de formation formelle (test de Student d'égalité des moyennes). Ce résultat confirme ceux obtenus par interrogation directe (Barron *et al.* [1997a] et [1997b] et Loewenstein et Spletzer [1994]). La formation formelle représente environ 22% de la formation totale. Cette part est supérieure à celle (environ 11%) obtenue par Barron *et al.* [1997a] et [1997b] à partir de mesures directes pour le premier mois d'ancienneté. La formation informelle étant dispensée essentiellement au cours des premières années d'ancienneté, il n'est pas surprenant que la part de la formation formelle dans la formation totale augmente dans le temps. Un test de Student sur le coefficient de corrélation linéaire entre formation formelle et formation informelle montre que ces deux types de formation sont corrélés positivement, ce qui suggère qu'elles sont complémentaires. Les établissements qui offrent le plus de formation formelle sont donc, en moyenne, ceux qui proposent également le plus de formation informelle. Notons que le choix des variables de contrôle, notamment des variables d'établissement (l'existence d'un système formalisé d'évaluation des performances individuelles; l'importance accordée à l'ancienneté du salarié, l'accroissement de ses performances, son effort de formation et son

accumulation d'expérience dans la progression individualisée du salaire ainsi que le nombre d'heures de formation formelle rémunérées par l'établissement au cours de l'année), effectué dans la partie III nous amène à penser que la formation que nous mesurons ici est bien informelle puisque ces variables suggèrent qu'une bonne partie des explications alternatives à la croissance des salaires avec l'ancienneté a été contrôlée. Par ailleurs, le fait que l'introduction ou non de variables de contrôle pour la formation formelle laisse le nombre d'heures de formation informelle quasiment inchangés (résultats non rapportés ici) signifie que la formation est essentiellement informelle. Cela étant, ce résultat demande à être confirmé dans la mesure où les variables de contrôle utilisées pour la formation formelle ne sont pas des variables individuelles mais des variables d'établissement.

CONCLUSION

L'utilisation d'un modèle de diffusion du savoir, qui identifie une des principales composantes de la formation en entreprise (le *learning by watching*), nous a permis de proposer une mesure économétrique de la formation informelle.

Ce modèle ne se limite pas à mesurer la formation par les autres, il fournit aussi une idée de la vitesse à laquelle se diffuse le savoir dans l'établissement et donne une estimation du savoir de l'établissement relativement à celui de l'individu au moment où ce dernier entre dans celui-ci.

Pour autant, aussi important soit-il, l'apprentissage des autres ne constitue pas le seul processus d'acquisition des connaissances en entreprise. Comme nous l'avons mentionné en introduction, il existe d'autres formes d'apprentissage informel. Malgré le nombre important de variables de contrôle utilisés ici, aucune d'entre elles ne prend en compte ces autres mécanismes d'apprentissage informel. Dès lors, la formation que nous mesurons ici peut ne pas être exclusivement du *learning by watching* mais l'ensemble de la formation informelle. Une des voies d'amélioration du modèle passe donc par l'introduction d'autres processus d'apprentissage dans les fonctions de gains.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ABOWD J.M. et KRAMARZ F. [1999]. "The analysis of Labor Markets Using Matched Employer-Employee Data", in *Handbook of Labor Economics*, vol.3, chap. 40, Ashenfelter O. and Card D. (eds).
- ABOWD J.M., KRAMARZ F., MARGOLIS D.N. et TROSKE K.R. [2000]. "Politiques salariales et performances des entreprises: une comparaison France/Etats-Unis", *Economie et Statistique*, 332-333 (2), pp. 27-38.
- ARAÏ M., BALLOT G. et SKALLI A. [1996]. "Rendements de l'ancienneté et taille des établissements", *Revue Economique*, 47 (3), pp. 623-632.
- BARRON J.M., BERGER M.C., et BLACK D.A. [1997a]. "How Well Do We Measure Training ?", *Journal of Labor Economics*, 15 (3), pp. 507-528.
- BARRON J.M., BERGER M.C., et BLACK D.A. [1997b]. *On-the-Job Training*, Kalamazoo, W.E. Upjohn Institute for Employment Research.
- BECKER G.S. [1962]. "Investment in Human Beings", *Journal of Political Economy*, 70 (5), pp. 9-49.
- BECKER G.S. [1975]. *Human Capital*, New-York, Columbia University Press for NBER (2^{ème} édition).
- BROWN C. [1989]. "Empirical Evidence On Private Training" in *Investing in People*, Background Papers, Vol. 1, Washington D.C.: Commission on Workforce Quality and Labor Market Efficiency, U.S. Department of Labor, pp. 301-320.
- CHENNOUF S., LEVY-GARBOUA L. et MONTMARQUETTE C. [1997]. "Les effets de l'appartenance à un groupe de travail sur les salaires individuels", *L'Actualité économique*, 73, pp. 207-232.
- DESTRE G. [2003]. "Fonctions de gains et diffusion du savoir: une estimation sur données françaises appariées", *Economie et Prévision* (à paraître).
- GOUX D. et MAURIN E. [1994]. "Education, expérience et salaire: tendances récentes et évolution de long terme", *Economie et Prévision*, 116 (5), pp. 155-178.
- LEVY-GARBOUA L. [1994]. "Formation sur le tas et rendements de l'expérience: un modèle de diffusion du savoir", *Economie et prévision*, 116 (5): 79-88.
- LOEWENSTEIN M.A. et SPLETZER J.R. [1994]. "Informal Training: A Review of Existing Data and Some New Evidence", Unpublished manuscript, Bureau of Labor Statistics.
- MINCER J. [1962]. "On-the-Job Training: Costs, Returns, and Some Implications", *Journal of Political Economy*, 70 (5), pp. 50-79.

MINCER J. [1974]. *Schooling, Experience, and Earnings*, New-York, Columbia University Press for NBER.

MINCER J. et JOVANOVIĆ B. [1981]. "Labor mobility and Wages", in S. Rosen (ed.), *Studies in labor Markets*, Chicago: Chicago University Press, pp. 21-64.

NORDMAN C. [2000]. " La formation sur le tas par diffusion du savoir: estimation sur données marocaines et mauriciennes ", *Revue d'Economie du Développement*, 4, pp. 79-103.

OCDE [1997]. *Manuel pour améliorer les statistiques de la formation*, Paris.

SICHERMAN N. [1990]. "The measurement of on-the-job Training", *Journal of Economic and Social Measurement*, 16 (4), pp. 221-230.

Annexe A
Tableau 1. Statistiques descriptives

<i>Symboles</i>	<i>Définition des variables</i>	Moyennes [Min ; Max] (Écart-types)
w_{ijt}	Gain horaire	71,81 [29,00 ; 394,73] (40,16)
nbh_{ran_i}	Nombre d'heures rémunérées dans l'année	1806,76 [33 ; 2310] (443,01)
$sexe_i$	1 pour les hommes, 0 pour les femmes	0,60
age_i	Age des salariés	38,84 [17 ; 65] (9,91)
nat_i	1 si français, 0 sinon ou non renseignée	0,93
$sitfam_i$	1 si marié, 0 sinon ou non renseignée	0,65
$cont_i$	1 si Contrat à durée indéterminée 0 sinon ou non renseignée	0,91
csp_i	1 si cadre, 0 sinon	0,12
s_i	Nombre d'années d'études	12,74 [8 ; 18] (1,66)
x_i	Nombre d'années d'expérience potentielle sur le marché du travail (hors de l'établissement)	9,10 [0 ; 47,66] (8,58)
t_{ij}	Nombre d'années d'ancienneté	10,63 [1 ; 46] (8,67)
N	Nombre d'observations	119 667
S_j	Sup de l'éducation	14,11 [9 ; 18] (1,92)
X_j	Sup de l'expérience potentielle hors de l'établissement	20,64 [0 ; 47,66] (9,34)
T_j	Sup de l'ancienneté	16,74 [1 ; 46] (10,20)
reg_j	1 pour Paris, 0 sinon	0,16
$effsal_j$	Effectif salarié total de l'établissement	141,81 (544,73)
$synd_j$	1 s'il y a des Délégués Syndicaux au sein de l'établissement, 0 sinon ou non renseignée	0,25
ep_j	1 s'il existe un système formalisé d'évaluation des performances individuelles au sein de l'établissement, 0 sinon ou non renseignée	0,15
$coop_j$	1 si les coopérations directes entre salariés sont encouragées, 0 sinon ou non renseignée	0,50
$rota_j$	1 si l'établissement pratique des rotations de poste, 0 sinon ou non renseignée	0,16
$evolsal_j$	1 si les évolutions de salaire de base dépendent exclusivement ou principalement des augmentations individualisées, 0 sinon ou non renseignée	0,05
$proganc_j$	1 si l'importance accordée à l'ancienneté du salarié dans la progression individualisée du salaire est très forte, 0 sinon ou non renseignée	0,06
$progperf_j$	1 si l'importance accordée à l'accroissement des performances du salarié dans la progression individualisée du salaire est très forte, 0 sinon ou non renseignée	0,35
$progform_j$	1 si l'importance accordée à l'effort de formation du salarié dans la progression individualisée du salaire est très forte, 0 sinon ou non renseignée	0,08
$progexp_j$	1 si l'importance accordée à l'accumulation d'expérience du salarié dans la progression individualisée du salaire est très forte, 0 sinon ou non renseignée	0,16
nbh_{fj}	Nombre d'heures moyen annuel de formation rémunérées par salarié dans l'établissement	11,92 [0 ; 1748,37] (54,01)
J	Nombre d'établissements	14 449

Tableau 2. Résultats des fonctions de gains (variable dépendante: \ln du gain horaire)

<i>Coefficients et paramètres</i>	FGM (Quadratique en t) (1)	LBW (FGB) (2)	LBW (FGN) (3)
<i>Coef. de s_i</i>	0,05363 ^a (0,00055)	0,06028 ^a (0,00061)	0,05912 ^a (0,00060)
<i>Coef. de x_i</i>	0,00951 ^a (0,00028)	0,01517 ^a (0,00039)	0,01453 ^a (0,00038)
<i>Coef. de x_i^2</i>	-0,00020 ^a (9,22 ^e -06)	-0,00033 ^a (0,00001)	-0,00032 ^a (0,00001)
<i>Coef. de t_{ij}</i>	0,01788 ^a (0,00030)		
<i>Coef. de t_{ij}^2</i>	-0,00017 ^a (9,78 ^e -06)		
β_s		0,44417 ^a (0,01565)	0,38550 ^a (0,01481)
β_x		0,90939 ^a (0,02526)	0,88750 ^a (0,02643)
<i>Coef. de T_{ij}</i>		0,02409 ^a (0,00080)	0,01966 ^a (0,00070)
<i>Coef. de T_{ij}^2</i>		-0,00041 ^a (0,00001)	-0,00034 ^a (0,00001)
n		0,04250 ^a (0,00195)	0,05494 ^a (0,00221)
r			0,41381 ^a (0,03036)
<i>Cste</i>	3,32723 ^a (0,01558)	3,22024 ^a (0,01597)	3,26508 ^a (0,01616)
<i>R² ajusté</i>	0,5977	0,5993	0,6000
N	119 667	119 667	119 667

Notes: Les écart-types à la moyenne figurent entre parenthèses

a: significatif à 1%

Tableau 3. Diffusion du savoir et savoir relatif¹⁵

	<i>n</i> (en %)	<i>t</i> _{0,50}	<i>H_{ij}/h_{i0}</i>
LBW (FGN)	5,49	12,95	1,40 (0,0004)
LBW (FGB)	4,25	16,64	1,50 (0,0005)

Tableau 4. Les rendements marginaux de l'éducation et de l'expérience initiale

	FGM (Quadratique en <i>t</i>)	LBW (FGB)	LBW (FGN)
$\frac{\partial \ln w_{ijt}}{\partial s_i}$	5,36	4,97	4,91
$\frac{\partial \ln w_{ijt}}{\partial x_i}$	0,57	0,57	0,51

Tableau 5. Les taux de rendement marginaux bruts de l'ancienneté (en %)

Modèles	FGM (Quadratique en <i>t</i>)	LBW (FGB)	LBW (FGN)
Première année d'ancienneté	1,78	2,06 ^a (0,002)	2,11 ^a (0,002)
Deuxième année d'ancienneté	1,75	1,93 ^a (0,002)	1,95 ^a (0,002)
Troisième année d'ancienneté	1,71	1,81 ^a (0,002)	1,81 ^a (0,002)
Quatrième année d'ancienneté	1,67	1,70 ^a (0,001)	1,68 ^a (0,001)
Cinquième année d'ancienneté	1,64	1,60 ^a (0,001)	1,56 ^a (0,001)
Ancienneté moyenne	1,44	1,16 ^a (0,001)	1,06 ^a (0,001)

¹⁵ Pour les tableaux 3 à 8, les chiffres entre parenthèses représentent les écart-types à la moyenne et a, b et c signifient respectivement significatif à 1, 5 et 10%

Tableau 6. *Nombre d'heures de formation informelle au cours de l'année 1992*

	Formation informelle (LBW)
Première année d'ancienneté	93,88 ^a (0,130)
Deuxième année d'ancienneté	86,84 ^a (0,119)
Troisième année d'ancienneté	80,48 ^a (0,109)
Quatrième année d'ancienneté	74,72 ^a (0,100)
Cinquième année d'ancienneté	69,47 ^a (0,092)
Ancienneté moyenne	47,38 ^a (0,060)

Tableau 7. *Comparaison entre le nombre d'heures moyen annuel de formation formelle et le nombre d'heures moyen annuel de formation informelle*

	Formation informelle	Formation formelle
e_{ijt} (moyenne par établissement)	42,75 ^a (0,150)	11,92 ^a (0,449)